GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 17:46:47 ; Search time 4900.12 Seconds
(without alignments)
16080.756 Million cell updates/sec

Title: US-10-018-786-6
Perfect score: 1818
Sequence: 1 atggcatacgcctgtcctcc.....cgcatatcgtctcgccctga 1818

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1818	100.0	1818	6	AX061805	AX061805 Sequence
2		1818	100.0	11629	1	AE012222	AE012222 Xanthomon
3		1203.4	66.2	11789	1	AE011667	AE011667 Xanthomon
4	c	1200.2	66.0	29206	1	AF499777	AF499777 Xanthomon
5		1196.6	65.8	1824	1	XANHRPA1A	M99173 Xanthomonas
6		1196.6	65.8	16929	1	XCU33548	U33548 Xanthomonas
7	c	1180.6	64.9	8230	1	AY139029S1	AY139029 Xanthomon
8	c	1180.6	64.9	19304	1	AB045311	AB045311 Xanthomon
9	c	1180.6	64.9	23514	1	AY205561	AY205561 Xanthomon
10	c	1180.6	64.9	100258	1	AY055110	AY055110 Xanthomon
11	c	1177.4	64.8	10096	1	AF026197	AF026197 Xanthomon
12		1175.8	64.7	2118	1	AB040401	AB040401 Xanthomon
13		467.2	25.7	1707	1	PSEHRP	M99631 Pseudomonas
14		467.2	25.7	23407	1	RSO245811	AJ245811 Ralstonia
15		467.2	25.7	197050	1	AL646081	AL646081 Ralstonia
16	c	358	19.7	29814	1	AF074878	AF074878 Burkholde
17		152.2	8.4	15231	1	AY166598	AY166598 Burkholde
18		152.2	8.4	34600	1	AY028431	AY028431 Burkholde
19		115.6	6.4	11557	1	AE004598	AE004598 Pseudomon
20		114	6.3	8424	1	PAU56077	U56077 Pseudomonas
21		100.6	5.5	344321	1	BX640429	BX640429 Bordetell
22		99	5.4	348624	1	BX640441	BX640441 Bordetell
23		97.6	5.4	1803	6	BD269173	BD269173 Vaccine.
24		97.6	5.4	1803	6	AX028410	AX028410 Sequence
25		97.6	5.4	35026	1	BPHYPLCRD	Y13383 Bordetella
26	c	97.6	5.4	348934	1	BX640417	BX640417 Bordetell
27		86.4	4.8	18931	1	AF501263	AF501263 Erwinia c
28		82.2	4.5	6176	1	EAU56662	U56662 Erwinia amy
29	c	81	4.5	204050	1	AL646073	AL646073 Ralstonia
30		79.2	4.4	125020	9	AF429315	AF429315 Homo sapi
31	c	75.4	4.1	47740	1	AY144116	AY144116 Photorhab
32		74	4.1	12037	1	PPUXPXCP	X81085 Pseudomonas
33		72.6	4.0	349944	1	BX571871	BX571871 Photorhab

	34	72.6	4.0	349980	6	AX770910	AX770910 Sequence
	35	72.2	4.0	20000	1	AF282857	AF282857 Pantoea s
c	36	71.8	3.9	125020	9	AF429315	AF429315 Homo sapi
	37	68.6	3.8	12148	15	AF302656	Af302656 Pectobact
c	38	68.6	3.8	28666	1	AY293288	AY293288 Pectobact
c	39	67.4	3.7	300917	1	AE016919	AE016919 Chromobac
	40	66.2	3.6	13320	1	AHEXE	X66504 A.hydrophil
	41	64.6	3.6	3923	1	PAXCPP	X68594 Pseudomonas
	42	64.6	3.6	11120	1	AE004734	AE004734 Pseudomon
c	43	63.8	3.5	16371	1	AF292566	AF292566 Pseudomon
	44	63.6	3.5	13197	1	AE004503	AE004503 Pseudomon
	45	63.4	3.5	11301	1	AE011699	AE011699 Xanthomon

ALIGNMENTS

RESULT 1
AX061805
LOCUS AX061805 1818 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 6 from Patent WO0078967.
ACCESSION AX061805
VERSION AX061805.1 GI:12539885
KEYWORDS .
SOURCE Xanthomonas campestris
ORGANISM Xanthomonas campestris
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xanthomonas.
REFERENCE 1
AUTHORS Pierrard,J., Simon,J.L. and Chevallereau,P.
TITLE Avirulent xanthomonas-campestris strains producing xanthan
JOURNAL Patent: WO 0078967-A 6 28-DEC-2000;
RHODIA CHIMIE (FR)
FEATURES Location/Qualifiers
source 1. .1818
/organism="Xanthomonas campestris"
/mol_type="unassigned DNA"
/db_xref="taxon:339"
ORIGIN

Query Match 100.0%; Score 1818; DB 6; Length 1818;
Best Local Similarity 100.0%; Pred. No. 3.3e-247;
Matches 1818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 60

Qy         61 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 120

Qy        121 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 180

Qy        181 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 240

Qy        241 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 300

Qy        301 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 360

Qy        361 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 420

Qy        421 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 480

Qy        481 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 540

Qy        541 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 600

Qy        601 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 660

Qy        661 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCGGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCGGG 720

Qy        721 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 780

Qy        781 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 840

Qy        841 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 900

Qy        901 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 960

Qy        961 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 1020

Qy       1021 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 1080

Qy       1081 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 1140

Qy       1141 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 1200

Qy       1201 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 1260

Qy       1261 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 1320

Qy       1321 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 1380

Qy       1381 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 1440

Qy       1441 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 1500

Qy       1501 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 1560

Qy       1561 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 1620

Qy       1621 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 1680

Qy       1681 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 1740

Qy       1741 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 1800

Qy       1801 CATATCGTCTCGCCCTGA 1818
             ||||||||||||||||||
Db       1801 CATATCGTCTCGCCCTGA 1818
```

RESULT 2
AE012222

```
LOCUS       AE012222               11629 bp    DNA     linear   BCT 23-MAY-2002
DEFINITION  Xanthomonas campestris pv. campestris str. ATCC 33913,  section 130
            of 460 of the complete genome.
ACCESSION   AE012222 AE008922
VERSION     AE012222.1  GI:21112273
KEYWORDS    .
SOURCE      Xanthomonas campestris pv. campestris str. ATCC 33913
  ORGANISM  Xanthomonas campestris pv. campestris str. ATCC 33913
            Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
            Xanthomonadaceae; Xanthomonas.
REFERENCE   1  (bases 1 to 11629)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Comparison of the genomes of two Xanthomonas pathogens with
            differing host specificities
  JOURNAL   Nature 417 (6887), 459-463 (2002)
  MEDLINE   22022145
   PUBMED   12024217
REFERENCE   2  (bases 1 to 11629)
  AUTHORS   da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
            Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
            Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
            Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
            Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
            Cursino-Santos,J.R., El-Dorry,H., Faria,J.B., Ferreira,A.J.S.,
            Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
            Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
            Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
            Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
            Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
            Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
            Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
            Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
            Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
            Kitajima,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade de
            Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
            Brazil
FEATURES             Location/Qualifiers
     source          1. .11629
                     /organism="Xanthomonas campestris pv. campestris str. ATCC
                     33913"
```

/mol_type="genomic DNA"
/strain="ATCC 33913"
/db_xref="ATCC:33913"
/db_xref="taxon:190485"
/note="pathovar: campestris"

gene complement(97. .708)
/gene="hpaP"
/note="synonym: XCC1228"

CDS complement(97. .708)
/gene="hpaP"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HpaP protein"
/protein_id="AAM40526.1"

/db_xref="GI:21112274"
/translation="MRILPIAGAPLQPHAAAAPARATQRVSFMQLRRRLSSVLAGLPD
TVVPEDGDDDPTLAEVMDAASDDHGPAALPTAPPCLDDGNRRQIARGDDGDALGRQIA
TEWIRTQRAQLAINHIALRVAEFCNATPVRGAGTWEAWLEINHDILPQTTLFLRLSPD
QLSLRFDSGSPETREVLLNGRKRLDTALKAALSDTLQISIEVV"

gene complement(732. .2654)
/gene="hrcV"
/note="synonym: XCC1229"

CDS complement(732. .2654)
/gene="hrcV"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrcV"
/protein_id="AAM40527.1"
/db_xref="GI:21112275"
/translation="MRVTRYFAYTGEVAIAALVVAVIGLMILPLPTPLIDTLLGINIT
LSVVLLMVTMYVPDSISLSSFPSLLLFTTLLRLSLNIASTKSILLHAEAGHIIESFGE
LVVGGNLVVGLVVFLIITTVQFIVIAKGSERVAEVGARFTLDAMPGKQMSIDADLRGG
NLTADEARRKRARLAIESQLHGGMDGAMKFVKGDAIAGLVITMVNILAGIVVGVTYHG
MSAGEAANRFAILSVGDAMVSQIASLLISVAAGVMITRVANENETKISSLGLDIGRQL
TSNARALMAASVLLACFAFVPGFPALLFLLLAAAVGAGGYTIWRKQRDTSGSDQPALP
STSRKGAKGDAPHIRKSAPDFASPLSMRLSPQLAARLDPALLDQAIESERRQLVELLG
LPFPGIAIWQSESLQGLQYEVLIHDVPETRSALSDTADMQKALAQQAIAPLHARAHLF
VGIQETQWMLEQVGADYPGLVAEVNKAMPAQRIADVLRRLLEERIPVRNIKSILESLV
VWGPKEKDLLMLTEYVRCDLGRYLAHTATAGTGQLPAVMLDHAVEQLIRQSIRATPAG
NFLALPPEQANQLVEQVERIVEDQARHPLAVVASMDVRRYVRRMIEARLNWLEVYSFQ
ELGAEVQLQPIGRVVA"

gene complement(2663. .3736)
/gene="hrcU"
/note="synonym: XCC1230"

CDS complement(2663. .3736)
/gene="hrcU"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrcU protein"
/protein_id="AAM40528.1"
/db_xref="GI:21112276"
/translation="MSDEKTEKPTEKKLQDARRDGEVPISPDVTAAAVLLAALLVMKL
AGSYFVEHLRALMSIGFDFTTNTRDATALHRALGRIGIQGVLLTLPFVTACLAAGLIG
TFVQTGLNASLKPVTPKFDSLNPVNGVKKLFSLRSLINLLKLGIKAAVIGVVLWYGLR
ALMPTIIGLAYQPPADIAQIGWRALGILCALAVLVFVLVGAADWSVQHWLFIRDKRMS
KDEQKREHKESEGDPEVKGKRKEFAKELVFGDPRERVAKAKVMVVNPTHYAVALAYEP
DGFGLPQVVAKGVDEGALELRAYAHNQGIPIVANPPLARALHEVELGEAVPESLFETV
AVVLRWVDELGRDNDEGSGPLPC"

gene 3949. .4404
/gene="hrpB1"
/note="synonym: XCC1231"

CDS 3949. .4404
/gene="hrpB1"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrpB1 protein"
/protein_id="AAM40529.1"
/db_xref="GI:21112277"
/translation="MEKIECPGSVVSGLIELITVGLTHEKIEEAAAVLAAVRVLRPEL
KALDTFDAWIAIKRGSYLEGARLLRELEADAGSEPLCKALYACCLFAVGDPSWHGIAE
GLIEEDADADAVGLVKALSGRYTPAPAPLEAAAESAAPMDVPNAQYLRA"

gene 4436. .4828
/gene="hrpB2"
/note="synonym: XCC1232"

CDS 4436. .4828
/gene="hrpB2"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrpB2 protein"
/protein_id="AAM40530.1"
/db_xref="GI:21112278"
/translation="MALIPPVVPITASTGLSTPAVSPTAAPNQSLVNRFQALMQSSCP
LPPAMQRVGEPSMISRMVDMQNDGVRAMTEHVDAFGANAPKMGLEEIAAEEMKLINEL
TMVGFNLNVSVSLAQSGKNAVQTLFKNQ"

gene 4830. .5594
/gene="hrcJ"
/note="synonym: XCC1233"

CDS 4830. .5594
/gene="hrcJ"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrcJ protein"
/protein_id="AAM40531.1"
/db_xref="GI:21112279"
/translation="MRTLRYLVVLLLALLLSGCDQQLYSGLTENDANDMLAVLLTAGV
DAEKLTPDDGKTWAVNAPHDQVAYALNVLRTHGMPHERHANLGEMFKKDGLISTPTEE
RVRFIYGVSQQLSQTLSNIDGVIAADVQIVLPNNDPLSASVKPSSAAVFIKFRVGSDL
TSLVPSIKTLVMHSVEGLTYENVSVTLVPGGAESDAQFAASAPPRAWAWPWLVGCALA
LCVAVAAAALYWWPSANARRWGGWQRLRALSRKHAG"

gene 5604. .6233
/gene="hrpB4"

/note="synonym: XCC1234"
CDS 5604. .6233
/gene="hrpB4"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrpB4 protein"
/protein_id="AAM40532.1"
/db_xref="GI:21112280"
/translation="MDKARIDNTARWLQRWTAAVSTLAQSLDAERVAHWLGPVAALGW
ERAAAQRRMRVIQAWSGWASMQVGALDPLANRLVVLAPDLLARVLMSRALFSRALALR
RCIERERLSWFEQRVGPAVLDYVRHCAGTGMTAPLLPHDADQAAWIGDGWRRMQADGV
WPDPCIAKVLALCLPLGAARVVPIAADGASDAFLRALPTLLPELPCVFG"
gene 6218. .6919
/gene="hrpB5"
/note="synonym: XCC1235"
CDS 6218. .6919
/gene="hrpB5"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="HrpB5 protein"
/protein_id="AAM40533.1"
/db_xref="GI:21112281"
/translation="MRLWLRSTPEAVGLDCEVIPREALACVLELDAAGAQVHARCAQA
LADAQTRAQALLDAAQRQAEAILQDAHDRAERSARLGYAAGLRRQLDAWNERGVRHAF
AAQDAARRARERLAEIVAHACEQVLHGHDPAALYARAAQALDGALDEANALQVSVHPD
ALDDARRAFDAAAAAGGWSMPVELCGDTTLALGACVCEWDTGVFETDLRDQLRSLRRV

Query Match 100.0%; Score 1818; DB 1; Length 11629;
Best Local Similarity 100.0%; Pred. No. 2.3e-247;
Matches 1818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 60
Db    9743 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 9802

Qy       61 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 120
Db    9803 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 9862

Qy      121 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 180
Db    9863 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 9922

Qy      181 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 240
Db    9923 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 9982

Qy      241 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 300
Db    9983 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 10042

Qy      301 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 360
```

```
Db   10043 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 10102

Qy      361 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 420
Db   10103 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 10162

Qy      421 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 480
Db   10163 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 10222

Qy      481 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 540
Db   10223 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 10282

Qy      541 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 600
Db   10283 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 10342

Qy      601 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 660
Db   10343 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 10402

Qy      661 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCCGG 720
Db   10403 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCGGG 10462

Qy      721 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 780
Db   10463 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 10522

Qy      781 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 840
Db   10523 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 10582

Qy      841 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 900
Db   10583 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 10642

Qy      901 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 960
Db   10643 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 10702

Qy      961 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 1020
Db   10703 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 10762

Qy     1021 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 1080
Db   10763 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 10822

Qy     1081 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 1140
Db   10823 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 10882

Qy     1141 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 1200
Db   10883 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 10942
```

```
Qy       1201 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10943 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 11002

Qy       1261 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11003 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 11062

Qy       1321 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11063 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 11122

Qy       1381 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11123 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 11182

Qy       1441 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11183 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 11242

Qy       1501 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11243 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 11302

Qy       1561 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11303 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 11362

Qy       1621 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11363 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 11422

Qy       1681 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11423 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 11482

Qy       1741 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11483 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 11542

Qy       1801 CATATCGTCTCGCCCTGA 1818
             ||||||||||||||||||
Db      11543 CATATCGTCTCGCCCTGA 11560
```

Search completed: July 7, 2004, 23:54:54
Job time : 4910.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 17:42:57 ; Search time 531.332 Seconds
(without alignments)
14535.578 Million cell updates/sec

Title: US-10-018-786-6
Perfect score: 1818
Sequence: 1 atggcatacgcctgtcctcc.....cgcatatcgtctcgccctga 1818

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1818	100.0	1818	5	ABL57894	Abl57894 Partial h
	2	97.6	5.4	1803	3	AAA64851	Aaa64851 Bordetell
	3	97.6	5.4	35026	3	AAA64890	Aaa64890 Bordetell
	4	72.6	4.0	1869	7	ACF71562	Acf71562 Photorhab
c	5	72.6	4.0	110000	7	ACF67367_47	Continuation (48 o
	6	72.6	4.0	110000	7	ACF65388_00	Acf65388 Photorhab
	7	68	3.7	114955	2	AAX53491	Aax53491 Human ade

	8	64.6	3.6	1977	7	ACA42395	Aca42395 Prokaryot
	9	62.2	3.4	2145	7	ABX94902	Abx94902 P. aurogi
	10	61.4	3.4	1752	7	ACA44288	Aca44288 Prokaryot
c	11	61	3.4	2751	7	ACA38169	Aca38169 Prokaryot
c	12	59	3.2	114955	2	AAX53491	Aax53491 Human ade
	13	56.8	3.1	2000	7	ADA71938	Ada71938 Rice gene
	14	56.8	3.1	2481	2	AAT13557	Aat13557 Syngp160m
	15	56.8	3.1	2481	2	AAT73952	Aat73952 HIV-1 gp1
	16	56.8	3.1	2481	2	AAV23291	Aav23291 Synthetic
	17	56.8	3.1	2571	2	AAZ08742	Aaz08742 HIV SYNgp
	18	56.8	3.1	2571	3	AAA93974	Aaa93974 Human imm
	19	56.8	3.1	2571	4	AAC86878	Aac86878 Nucleotid
	20	56.8	3.1	2571	5	AAH43701	Aah43701 Codon opt
	21	56.6	3.1	1200	7	ACA26511	Aca26511 Prokaryot
c	22	55.4	3.0	65140	4	AAD17184	Aad17184 Streptomy
c	23	55.4	3.0	125401	4	AAD17186	Aad17186 Streptomy
	24	55.2	3.0	12588	2	AAQ63293	Aaq63293 Sequence
	25	54.8	3.0	103599	4	ABX04971	Abx04971 S. cinnam
c	26	54.4	3.0	30365	4	AAD20405	Aad20405 P. syring
	27	54.2	3.0	1746	7	ACA27403	Aca27403 Prokaryot
	28	54	3.0	1162	3	AAX89451	Aax89451 Rice bran
c	29	53.8	3.0	88421	6	AAL40781	Aal40781 88421nt g
	30	53.6	2.9	1227	7	ADA70473	Ada70473 Rice gene
c	31	53	2.9	2000	7	ADA71938	Ada71938 Rice gene
c	32	52.8	2.9	113193	7	AAD54645	Aad54645 Streptomy
	33	52.8	2.9	349980	6	ABQ81849	Abq81849 Bifidobac
	34	52.4	2.9	536	9	ADB68842	Adb68842 Minority
	35	52.4	2.9	110000	4	AAI99683_08	Continuation (9 of
c	36	52	2.9	16020	3	AAA39283	Aaa39283 Streptomy
	37	51.6	2.8	2331	4	AAS51455	Aas51455 Pseudomon
	38	51.6	2.8	2331	7	ACA19557	Aca19557 Prokaryot
	39	51.6	2.8	2772	7	ACA40420	Aca40420 Prokaryot
	40	51.6	2.8	15672	2	AAQ10613	Aaq10613 Rianodin
	41	51.6	2.8	110000	4	AAI99682_10	Continuation (11 o
	42	51.6	2.8	110000	4	AAI99683_10	Continuation (11 o
c	43	51.2	2.8	34720	2	AAV30458_5	Continuation (6 of
c	44	51.2	2.8	36165	2	AAV30459_5	Continuation (6 of
c	45	51.2	2.8	110000	2	AAV30456_4	Continuation (5 of

ALIGNMENTS

RESULT 1
ABL57894
ID ABL57894 standard; DNA; 1818 BP.
XX
AC ABL57894;
XX
DT 11-SEP-2003 (revised)
DT 04-JUL-2002 (first entry)
XX
DE Partial hypersensitive reaction and pathogenicity, hrpA gene.
XX
KW Hypersensitive reaction and pathogenicity; hrpA; exo-polysaccharide;
KW xanthan gum; gene; ds.
XX
OS Xanthomonas campestris; pv vesicatoria.
XX
PN WO200078967-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-FR001725.
XX
PR 22-JUN-1999; 99FR-00007963.
XX
PA (RHOD) RHODIA CHIM.
XX
PI Pierrard J, Simon J, Chevallereau P;
XX
DR WPI; 2001-102725/11.
XX
PT New Xanthomonas campestris bacteria strains for use in production of exo-
PT polysaccharides are made non-virulent by inactivation of at least one
PT virulence gene.
XX
PS Claim 18; Page 26-27; 33pp; French.
XX
CC The present invention relates to new Xanthomonas campestris bacteria
CC strains made non-virulent by inactivation of at least one virulence gene
CC but which have retained the capacity to produce exo-polysaccharides
CC (preferably xanthan gum). One such virulence gene deleted to produce the
CC bacterial strains was the hrpA gene (Hypersensitive Reaction and
CC Pathogenicity). The hrp genes are essential for pathogenicity in plants.
CC The present sequence is a partial sequence of the hrpA, used in an
CC example from the invention. (Updated on 11-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 1818 BP; 306 A; 582 C; 602 G; 328 T; 0 U; 0 Other;

Query Match 100.0%; Score 1818; DB 5; Length 1818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCATACGCCTGTCCTCCAGTTCACCGCCATCGACGCGCGCCGTTGGCCGCTGCCTTG 60

Qy         61 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGCTTGGCTTGCTGCCGCTGCTGCCGCCGCATGCCAACGCCGCGTCGGTGCCGTGGCAC 120

Qy        121 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGCGCAGCTTCAAATACGTTGCCGACCGCAAGGATCTCAAGGAGGTGCTGCGCGACCTG 180

Qy        181 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGCCAGCCAATCCATCACCACCTGGATTTCACCGGAGGTGACCGGCACGCTCAGTGGC 240

Qy        241 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAATTCGAAGCCACTCCGCAGAAGTTTCTCGACGATCTATCGGGCACGTTCGGTTTTGTC 300
```

```
Qy       301 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TGGTATTACGATGGCTCGGTGCTCAGAATCTGGGGCGCGAACGAGACCAAGAATGCGACC 360

Qy       361 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TTGAGTTTGGGCGCTGCATCGACGAGTGCGCTGCGCGATGCGCTTGCGCGCATGCGGCTG 420

Qy       421 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GACGATCCGCGCTTTCCGGTCCGTTATGACGAGACAGCGCACCTGGCGGTGGTGTCGGGC 480

Qy       481 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CCGCCGGGTTATGTGGATACCGTCGCGGCGATCGCCAAGCAGGTCGAGCAGGTCGCGCGC 540

Qy       541 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CAACGCGACGCCACCGAAGTGCAGGTGTTTCAGCTGCATTATGCGCAGGCGGCCGACCAC 600

Qy       601 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ACCACCCGCATCGGTGGTCAAGACATCCAGGTGCCGGGCATGGCCAGCCTGTTGCGCAAC 660

Qy       661 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCGGG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 ATATACGGCGTGCGTGGCGCGCCCACTGCGGCGCTGCCCGGGCCAGGCGCGAATTTCGGG 720

Qy       721 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CGTGTGCAACCGATCGGCGGTGGCTCGTCCAATACCTTCGGCAACAGCGGTCAGCGCCAG 780

Qy       781 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 AGTGGCGGCAGCGGCATTCTCGGTTTGCCTGCGTCGTGGTTCGGCGCTGGGTCGCCGTCC 840

Qy       841 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GAGCGGGTGCCGGTCAGTCCGCCGTTGCCGGGCAGTGGCAATAGCGCCAATGCGCCGGCC 900

Qy       901 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AGCGTGTGGCCGGAGATGAGCCAGGCCAGACGCGATGCGCCGCTGGCGGTGGACGCCGGC 960

Qy       961 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 AGCGGCGGTGAGCTGGCATCCGACGCGCCGGTGATCGAAGCCGACCCGCGCACCAACGGC 1020

Qy      1021 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 ATTCTCATTCGCGACCGCCCCGAGCGGATGGCCGCCTATGGCACGTTGATCCAGCAGCTC 1080

Qy      1081 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 GACAACCGTCCCAAGCTGCTGCAGATCGATGCCACCATCATCGAGATCCGCGACGGCGCC 1140

Qy      1141 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CTGCAGGATCTCGGCGTGGACTGGCGGTTCCACAGCCGGCGTGTGGATGTGCAGACCGGC 1200

Qy      1201 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GACGGGCGTGGTGGCCAGCTTGGCTACGATGGCAGCTTGAGCGGTGCAGCAGCCGCCGGT 1260

Qy      1261 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 GCAGCCGCGCCGTTGGGCGGGACGTTGACCGCTGTCCTGGGCGATGCAGGGCGTTACCTG 1320

Qy      1321 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ATGACGCGCGTCTCGGCGCTCGAGCAGACCAACAAGGCCAAGATCGTCTCCACCCCGCAG 1380

Qy      1381 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 GTGGCGACGCTGGACAACGTGGAAGCGGTGATGGACCACAAGCAACAGGCATTCGTGCGT 1440

Qy      1441 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GTCAGCGGTTATGCATCCGCCGACCTCTACAACCTGTCCGCGGGTGTATCGCTACGCGTA 1500

Qy      1501 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 TTGCCAAGTGTGGTGCCGGGGTCGCCAAATGGTCAGATGCGCCTGGATGTGCGTATCGAA 1560

Qy      1561 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GACGGGCAGTTGGGCGCCAATACCGTCGATGGCATTCCCGTCATCACCTCCAGCGAGATC 1620

Qy      1621 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 ACCACGCAGGCCTTCGTCAACGAGGGCCAGAGCCTGCTGATCGCCGGTTATGCTTCCGAC 1680

Qy      1681 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 ACCGATCAGACAGATCTGAACAACGTCCCCGGGCTGTCCAGGATTCCATTGGTCGGCAAC 1740

Qy      1741 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 CTGTTCAAGCATCGCCAGCAGAGCGGGTCGCGGTTGCAGCGGTTGTTCCTGCTGACCCCG 1800

Qy      1801 CATATCGTCTCGCCCTGA 1818
             ||||||||||||||||||
Db      1801 CATATCGTCTCGCCCTGA 1818
```

Search completed: July 7, 2004, 21:07:49
Job time : 537.332 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 20:18:17 ; Search time 3500.23 Seconds
(without alignments)
15510.269 Million cell updates/sec

Title: US-10-018-786-6
Perfect score: 1818
Sequence: 1 atggcatacgcctgtcctcc.....cgcatatcgtctcgccctga 1818

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	74.2	4.1	719	28	BZ559988	BZ559988 pacs2-164
c	2	70	3.9	932	29	CNS0072Q	AL066742 Drosophil
c	3	66.6	3.7	1507	28	BH614483	BH614483 2B8_Conti
c	4	66.6	3.7	2481	28	BH614291	BH614291 1C22/2B8_
c	5	65.4	3.6	925	29	CNS0091P	AL053013 Drosophil
c	6	63	3.5	1101	29	CNS017SY	AL108460 Drosophil
c	7	62.2	3.4	889	28	BZ571081	BZ571081 msh2_172.
	8	62.2	3.4	935	29	CNS006XK	AL066051 Drosophil
	9	61.6	3.4	925	29	CNS0091P	AL053013 Drosophil
c	10	61	3.4	589	14	CB852504	CB852504 UI-CF-FN0
c	11	61	3.4	1004	14	CD051228	CD051228 AGENCOURT
c	12	61	3.4	1288	13	BQ678719	BQ678719 AGENCOURT
c	13	60.4	3.3	1151	14	CK210749	CK210749 FGAS02257
	14	60	3.3	982	13	BX415111	BX415111 BX415111
c	15	59.6	3.3	723	14	CB854525	CB854525 UI-CF-DU1
c	16	59.6	3.3	781	14	CB851140	CB851140 UI-CF-DU1
c	17	59.6	3.3	982	13	BX415111	BX415111 BX415111
c	18	59.6	3.3	1085	28	BZ552651	BZ552651 pacs1-60_
c	19	59.4	3.3	935	29	CNS006XK	AL066051 Drosophil
c	20	59.2	3.3	1201	13	BX381961	BX381961 BX381961
c	21	59.2	3.3	1625	29	AG043477	AG043477 Pan trogl
	22	58.4	3.2	1421	28	BZ569488	BZ569488 pacs2-164
	23	58.2	3.2	740	29	CNS01QUO	AL156592 Anopheles
	24	58.2	3.2	1250	29	AG043469	AG043469 Pan trogl
c	25	58	3.2	888	29	AG030591	AG030591 Pan trogl
c	26	57.8	3.2	928	29	AG054567	AG054567 Pan trogl
	27	57.8	3.2	1452	29	AG032979	AG032979 Pan trogl
c	28	57.2	3.1	1311	12	BM544820	BM544820 AGENCOURT
	29	56.6	3.1	921	28	AZ528613	AZ528613 ENTCM21TR
	30	56.6	3.1	932	29	CNS0072Q	AL066742 Drosophil
c	31	56.2	3.1	1150	13	BQ892481	BQ892481 AGENCOURT
	32	56.2	3.1	1278	29	AG060116	AG060116 Pan trogl
c	33	56.2	3.1	1281	12	BG852363	BG852363 1024034A0
	34	56.2	3.1	1542	29	AG032943	AG032943 Pan trogl
	35	56	3.1	1516	12	BG809984	BG809984 mgct002xd
c	36	56	3.1	1610	28	BZ569386	BZ569386 pacs2-164
	37	55.8	3.1	1057	13	BX349688	BX349688 BX349688
c	38	55.8	3.1	1339	12	BM458211	BM458211 AGENCOURT
c	39	55.8	3.1	1598	29	AG030579	AG030579 Pan trogl
c	40	55.4	3.0	927	13	BQ680645	BQ680645 AGENCOURT
	41	55.4	3.0	1201	13	BX356664	BX356664 BX356664
c	42	55.4	3.0	1798	29	AG171124	AG171124 Pan trogl
	43	55.2	3.0	848	29	AG036341	AG036341 Pan trogl
	44	55.2	3.0	895	29	AG064040	AG064040 Pan trogl
c	45	55.2	3.0	1476	29	AG058501	AG058501 Pan trogl

Search completed: July 8, 2004, 01:54:08
Job time : 3505.23 secs